

Journal of Econometrics 19 (1982) 7-22. North-Holland Publishing Company

A BAYESIAN APPROACH TO RETROSPECTIVE IDENTIFICATION OF CHANGE-POINTS*

N. B. BOOTH

Polytechnic of North London, London, UK

A.F.M. SMITH

University of Nottingham, Nottingham NG7 2RD, UK

1. Introduction

We begin our discussion of a general Bayesian approach to making inferences about change-points by considering a simple special case.

Suppose that y_1, \dots, y_n are realizations of a random process such that, in the absence of any change in underlying structure, the joint density of y_1, \dots, y_n would have the form

$$p(y_1, \dots, y_n | \psi_1) = \prod_{i=1}^n p(y_i | \psi_1), \quad (1)$$

where ψ_1 is a vector of parameters. In such a case, we say that the process exhibits a change-point at time r ($1 \leq r < n$) if there exists ψ_2 ($\psi_2 \neq \psi_1$) such that

$$\begin{aligned} p(y_1, \dots, y_n | r, \psi_1, \psi_2) &= p(y_1, \dots, y_r | \psi_1) p(y_{r+1}, \dots, y_n | \psi_2) \\ &= \prod_{i=1}^r p(y_i | \psi_1) \prod_{i=r+1}^n p(y_i | \psi_2). \end{aligned} \quad (2)$$

[The notation $p(\cdot | \cdot)$ is simply a generic symbol for a conditional density; its actual interpretation will always be clear from the context.]

We shall denote by M_r the model which assumes there to be a change point at r . In the case of known ψ_1, ψ_2 , we therefore have

$$p(y_1, \dots, y_n | M_r) = p(y_1, \dots, y_n | r, \psi_1, \psi_2), \quad (3)$$

*The material in section 4 is based on Chapter 2 of the first author's Ph.D. thesis written at University College, London, under the supervision of the second author.

given by (2), whereas, if ψ_1, ψ_2 are unknown, we have

$$p(y_1, \dots, y_n | M_r) = \iint p(y_1, \dots, y_n | r, \psi_1, \psi_2) p(\psi_1, \psi_2) d\psi_1 d\psi_2, \quad (4)$$

where $p(\psi_1, \psi_2)$ specifies a prior density for ψ_1, ψ_2 (here assumed independent of M_r). It is convenient to denote by M_0 the model which assumes that no change in underlying distribution has occurred. In the case of unknown ψ_1 , we therefore have

$$p(y_1, \dots, y_n | M_0) = \int p(y_1, \dots, y_n | \psi_1) p(\psi_1) d\psi_1, \quad (5)$$

whereas, in the case of known ψ_1 , the left-hand side of (5) is given directly by eq. (1).

Given y_1, \dots, y_n inferences about the change-point are equivalent to inferences about the alternative models M_0, M_1, \dots, M_{n-1} . The basis for such inferences is provided straightforwardly by Bayes' theorem, which takes the form

$$p(M_r | y_1, \dots, y_n) \propto p(y_1, \dots, y_n | M_r) p(M_r), \quad (6)$$

where $p(M_r)$ denotes the prior probability of a change-point occurring at r .

Thus, for example, comparison of the posterior probabilities of two alternative hypotheses M_r, M_s , say, is given by the ratio of the two terms (for r and s , respectively) provided by (6). A convenient way of re-expressing this is in the form

$$\frac{p(M_r | y_1, \dots, y_n) / p(M_r)}{p(M_s | y_1, \dots, y_n) / p(M_s)} = \frac{p(y_1, \dots, y_n | M_r)}{p(y_1, \dots, y_n | M_s)}, \quad (7)$$

displaying the ratio of posterior to prior odds on M_r against M_s as a (usually, integrated) likelihood ratio. The right-hand side of (7) is called the Bayes factor for M_r against M_s and will be denoted by B_{rs} . A value of $B_{rs} > 1$ thus indicates evidence from the data pointing towards r as a more likely change-point than s . In particular, if $s=0$, B_{r0} provides an indicator of whether a change-point at r is better supported than an assumption of no change. Since $B_{0r} = B_{r0}^{-1}$ and $B_{rs} = B_{r0} B_{0s}$, it will suffice, in what follows, to work in terms of B_{r0} .

If we are interested in an overall assessment of change versus no change, we note that the appropriate ratio of posterior to prior odds is given by

$$\left(\frac{1 - p(M_0 | y_1, \dots, y_n)}{p(M_0 | y_1, \dots, y_n)} \right) / \left(\frac{1 - p(M_0)}{p(M_0)} \right) = \sum_r B_{r0} \left(\frac{p(M_r)}{1 - p(M_0)} \right). \quad (8)$$

The right-hand side Bayes factors for probabilities of those particular, if the P which they are non-taken over all the p

We began our dis (1)–(5). This will be means in uni- and blems of changes in and 4), (1)–(5) no $p(y_1, \dots, y_n | M_r)$ occur inferences about r within the Bayesian unknown parameter

Detailed Bayesian been given in Brox (1977) and Menzefri briefly in section 2 regression and tim previously on these Smith (1980), but w ideas in two new ir to the problem of assumes change, ur parameters within different dimension assigning values to specification of forr overcoming the pr provide a unified treatment of the tin technique used is likelihood for ARM

Throughout, we change. For Bayesi Harrison and Steve (1982).

2. Changes of mean

2.1. The univariate

Bayesian inferen- sequence is dealt w

$$\psi_1, \psi_2) d\psi_1 d\psi_2, \quad (4)$$

here assumed independent model which assumes that In the case of unknown

$$\psi_1, \quad (5)$$

of (5) is given directly by

nt are equivalent to in-
 M_{n-1} . The basis for such
 theorem, which takes the

(6)

point occurring at r .
 or probabilities of two
 atio of the two terms (for
 way of re-expressing this

(7)

in M_r against M_0 as a
 side of (7) is called the
 by B_{r0} . A value of $B_{r0} > 1$
 ards r as a more likely
 rovides an indicator of
 an an assumption of no
 fice, in what follows, to

change versus no change,
 or odds is given by

$$B_{r0} = \left(\frac{P(M_r)}{1 - P(M_0)} \right). \quad (8)$$

The right-hand side of (8) has the form of a weighted average of individual Bayes factors for changes at specific times, weighted by the a priori probabilities of those changes, conditional on a change having occurred. In particular, if the $P(M_r)/(1 - P(M_0))$ terms are equal over the range of r for which they are non-zero, then (8) is just an average of the Bayes factors B_{r0} , taken over all the possible change-points.

We began our discussion by considering a simple structure, summarized in (1)–(5). This will be illustrated in section 2, where we consider changes of means in uni- and multivariate normal sequences. For more general problems of changes in structure of regression or time series models (sections 3 and 4), (1)–(5) no longer apply, but, provided we can specify the form $p(y_1, \dots, y_n | M_r)$ occurring in (6), the forms (6)–(8) still provide the basis for inferences about r . In addition to inferences about r , it is straightforward within the Bayesian approach to provide inference statements about other unknown parameters occurring in the specification of model structures.

Detailed Bayesian analysis of parameter shifts in univariate series have been given in Broemeling (1972, 1974), Smith (1975), Lee and Heghinian (1977) and Menzefricke (1981). Some aspects of this work will be summarized briefly in section 2, but in this paper we shall concentrate mainly on the regression and time series cases (sections 3 and 4). We have reported previously on these problems in Smith (1977), Smith and Cook (1980) and Smith (1980), but we shall be concerned here with developing our previous ideas in two new important directions. First, we provide a possible solution to the problem of comparing the no-change model with a model which assumes change, under the assumption of vague prior information for the parameters within each model. Since the models being compared have different dimensions, we are forced to confront directly the problem of assigning values to the (usually ignored) arbitrary constants arising in the specification of forms of improper prior density. Our basic suggestion for overcoming the problem will be introduced in section 2. Secondly, we provide a unified account of a procedure which greatly facilitates the treatment of the time series case by relating it to a regression problem. The technique used is closely related to the problem of finding the exact likelihood for ARMA processes.

Throughout, we shall be concerned with retrospective inference about change. For Bayesian approaches to on-line monitoring, see, for example, Harrison and Stevens (1976), Smith and Makov (1980) and Smith and West (1982).

2. Changes of mean in normal sequences

2.1. The univariate case

Bayesian inference about a change of mean in a univariate normal sequence is dealt with in detail by Smith (1975, sec. 4), and the change of

variance case is treated by Menzefricke (1981). In this section, we shall just consider the problem of comparing models of change of mean versus no change when there is vague prior knowledge about the model parameters. We begin with the simple case of a conditionally independent univariate normal sequence and then generalize to the multivariate case. This will enable us to highlight the problems which arise in calculating B_{r0} , and hence (8), when standard improper prior forms of representation of vague knowledge are assumed.

Corresponding to (2), with $\psi_j = (\mu_j, \sigma)$, we shall assume that, for M_r with $r \neq 0$,

$$p(y_1, \dots, y_n | r, \psi_1, \psi_2) \\ = (2\pi\sigma^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \left[\sum_{i=1}^r (y_i - \mu_1)^2 + \sum_{i=r+1}^n (y_i - \mu_2)^2 \right] \right\}.$$

Further, we shall assume a prior density $p(\psi_1, \psi_2)$ of the form

$$p(\mu_1, \mu_2, \sigma) = p(\mu_1 | \sigma) p(\mu_2 | \sigma) p(\sigma) \\ = \{c_1(2\pi\sigma^2)^{-1/2}\} \{c_2(2\pi\sigma^2)^{-1/2}\} c\sigma^{-1}, \quad (9)$$

where c_1, c_2, c are unspecified constants [see, for example, Jeffreys (1961)].

It then follows from (4), after some straightforward algebra, that

$$p(y_1, \dots, y_n | M_r) \\ \propto c_1 c_2 c [r(n-r)]^{-1/2} \left[\sum_{i=1}^r (y_i - \bar{y}_r)^2 + \sum_{i=r+1}^n (y_i - \bar{y}_{(n-r)})^2 \right]^{-n/2}, \quad (10)$$

where $\bar{y}_r, \bar{y}_{(n-r)}$ denote the mean of y_1, \dots, y_r and y_{r+1}, \dots, y_n , respectively. In the case of M_0 , we take $p(\psi_1) = p(\mu_1 | \sigma) p(\sigma)$ and obtain

$$p(y_1, \dots, y_n | M_0) \propto c_1 c n^{-1/2} \left[\sum_{i=1}^n (y_i - \bar{y}_n)^2 \right]^{-n/2}. \quad (11)$$

We now note that if we take the ratio of (10) and (11) in order to obtain B_{r0} from (7), we obtain, after some algebra,

$$B_{r0} = c_2 \left(\frac{n}{r(n-r)} \right)^{1/2} \left(1 + \frac{t_r^2}{n-2} \right)^{n/2}, \quad (12)$$

where t_r^2 is the usual t sampling distribution d connection between B comparing means. In f minimum (i.e., showir (corresponding to $\bar{y}_r = \bar{y}$

However, the unple depends on the unspe arise if we calculate B defined by (10), the fa and M_0 , assuming vag different dimensionality under M_r) and arbitra implicitly, from impro (1971) and Atkinson (1

In order to motivate a 'thought experiment Spiegelhalter (1981) an

We imagine that we to compare M_0 and M in favour of M_r . This such a case, we should we have only one obs much from 1. If we ent represent our unwilling very minimal evidence, choice of c_2 .

In fact, we have as should not depend on sample' in order for B , corresponding to $t_r^2 =$ appropriate choice for $n = r + 1 = 3$, which ju unknown parameters $B_{20} = 1$, into (12) we ot

2.2. The multivariate ca

Suppose that under

$$y_i \sim N_p(\mu_1,$$

$$y_i \sim N_p(\mu_2,$$

variance case is treated by Menzefricke (1981). In this section, we shall just consider the problem of comparing models of change of mean versus no change when there is vague prior knowledge about the model parameters. We begin with the simple case of a conditionally independent univariate normal sequence and then generalize to the multivariate case. This will enable us to highlight the problems which arise in calculating $B_{r,0}$, and hence (8), when standard improper prior forms of representation of vague knowledge are assumed.

Corresponding to (2), with $\psi_j = (\mu_j, \sigma)$, we shall assume that, for M_r with $r \neq 0$,

$$p(y_1, \dots, y_n | r, \psi_1, \psi_2) \\ = (2\pi\sigma^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \left[\sum_{i=1}^r (y_i - \mu_1)^2 + \sum_{i=r+1}^n (y_i - \mu_2)^2 \right] \right\}.$$

Further, we shall assume a prior density $p(\psi_1, \psi_2)$ of the form

$$p(\mu_1, \mu_2, \sigma) = p(\mu_1 | \sigma) p(\mu_2 | \sigma) p(\sigma) \\ = \{c_1(2\pi\sigma^2)^{-1/2}\} \{c_2(2\pi\sigma^2)^{-1/2}\} c\sigma^{-1}, \quad (9)$$

where c_1, c_2, c are unspecified constants [see, for example, Jeffreys (1961)].

It then follows from (4), after some straightforward algebra, that

$$p(y_1, \dots, y_n | M_r) \\ \propto c_1 c_2 c [r(n-r)]^{-1/2} \left[\sum_{i=1}^r (y_i - \bar{y}_r)^2 + \sum_{i=r+1}^n (y_i - \bar{y}_{(n-r)})^2 \right]^{-n/2}, \quad (10)$$

where $\bar{y}_r, \bar{y}_{(n-r)}$ denote the mean of y_1, \dots, y_r and y_{r+1}, \dots, y_n , respectively. In the case of M_0 , we take $p(\psi_1) = p(\mu_1 | \sigma) p(\sigma)$ and obtain

$$p(y_1, \dots, y_n | M_0) \propto c_1 c n^{-1/2} \left[\sum_{i=1}^n (y_i - \bar{y}_n)^2 \right]^{-n/2}. \quad (11)$$

We now note that if we take the ratio of (10) and (11) in order to obtain $B_{r,0}$ from (7), we obtain, after some algebra,

$$B_{r,0} = c_2 \left(\frac{n}{r(n-r)} \right)^{1/2} \left(1 + \frac{t_r^2}{n-2} \right)^{n/2}, \quad (12)$$

where t_r^2 is the usual sampling distribution connection between comparing means. The minimum (i.e., the χ^2 corresponding to \bar{y} ,

However, the un depends on the un arise if we calculate defined by (10), the and M_0 , assuming different dimension: under M_r) and arise implicitly, from imp (1971) and Atkinson

In order to motivate a 'thought experiment' Spiegelhalter (1981)

We imagine that to compare M_0 and in favour of M_r . The such a case, we show we have only one choice much from 1. If we represent our unwilling very minimal evidence choice of c_2 .

In fact, we have should not depend sample' in order for corresponding to the appropriate choice: $n = r + 1 = 3$, which unknown parameter $B_{2,0} = 1$, into (12) we

2.2. The multivariate

Suppose that under

$$y_i \sim N_p(\mu_1, \Sigma)$$

$$y_i \sim N_p(\mu_2, \Sigma)$$

$$\left. \begin{aligned} y_i &\sim N_p(\mu_1, \Sigma), & i=1, \dots, r \\ y_i &\sim N_p(\mu_2, \Sigma), & i=r+1, \dots, n \end{aligned} \right\}, \quad (13)$$

where the observation y_i is now a p -vector, and that under M_0 we assume

$$y_i \sim N_p(\mu_1, \Sigma), \quad i = 1, \dots, n. \quad (14)$$

In both cases, the sequences of observations are assumed conditionally independent. For the unknown parameters μ_1, μ_2, Σ , invariance arguments [Jeffreys (1961)] lead to improper prior specifications of the form

$$p(\mu_1, \mu_2, \Sigma) = p(\mu_1 | \Sigma) p(\mu_2 | \Sigma) p(\Sigma), \quad (15)$$

where

$$\left. \begin{aligned} p(\mu_j | \Sigma) &= c_j (2\pi)^{-p/2} |\Sigma|^{-1/2} \\ p(\Sigma) &= c |\Sigma|^{-(p+1)/2} \end{aligned} \right\}. \quad (16)$$

Using standard integration formulae [see, for example, Smith and Spiegelhalter (1981, sec. 17.2.2)], we can easily find the forms of

$$\begin{aligned} p(y_1, \dots, y_n | M_r) &= \iint p(y_1, \dots, y_n | \mu_1, \mu_2, \Sigma) \\ &\quad \times p(\mu_1 | \Sigma) p(\mu_2 | \Sigma) p(\Sigma) d\mu_1 d\mu_2 d\Sigma, \end{aligned} \quad (17)$$

and

$$p(y_1, \dots, y_n | M_0) = \iint p(y_1, \dots, y_n | \mu_1, \Sigma) p(\mu_1 | \Sigma) p(\Sigma) d\mu_1 d\Sigma. \quad (18)$$

Taking the ratio of (17) and (18), we find that

$$B_{r0} = c_2 \left(\frac{n}{r(n-r)} \right)^{p/2} \left(\frac{|S_n|}{|S_r + S_{(n-r)}|} \right)^{n/2}, \quad (19)$$

where

$$\left. \begin{aligned} S_r &= \sum_{i=1}^r (y_i - \bar{y}_r)(y_i - \bar{y}_r)^T \\ S_{(n-r)} &= \sum_{i=r+1}^n (y_i - \bar{y}_{(n-r)})(y_i - \bar{y}_{(n-r)})^T \end{aligned} \right\}. \quad (20)$$

Arguing as in the previous section, we see that a minimal calibrating sample is provided by $r=p$, $n=2p+1$. If this gave $\bar{y}_r = \bar{y}_{(n-r)}$ so that $S_r + S_{(n-r)} = S_n$, we should want, approximately, $B_{r0} = 1$. Substitution into (19) leads to the choice

$$c_2 = \left(\frac{p(p+1)}{2p+1} \right)^{p/2}. \quad (21)$$

Using this value, (1) spectve inferences : normal observations riance structure can Smith and Spiegelha

3. Changes of coeffi

3.1. Introduction

We shall consider

$$Y_t = x_t^T \beta^{(t)}$$

where, at time t , Y_t column vector of ob constant term), $\beta^{(t)}$ i and ε_t is a disturban differing t , with zero

The sequence def $(1 \leq r < n)$ if

$$\beta^{(1)} = \dots$$

with unknown β , δ change, $\delta=0$, will be

If we write $Y = ($ adopting the notatio

$$X_r^T = (x_1,$$

we see that, under Λ

$$Y \sim N_n(A$$

where

$$A_r = \begin{pmatrix} X \\ X_{(n)} \end{pmatrix}$$

assume

(14)

defined conditionally
invariance arguments
of the form

(15)

(16)

example, Smith and Spiegel-
mans of

)

$d\mu_1 d\mu_2 d\Sigma$, (17)

$1 | \Sigma) p(\Sigma) d\mu_1 d\Sigma$. (18)

(19)

(20)

at a minimal calibrating
gave $\bar{y}_r = \bar{y}_{(n-r)}$ so that
= 1. Substitution into (19)

(21)

Using this value, (19), in conjunction with (8), provides a basis for retrospective inferences about changes of mean in a sequence of multivariate normal observations. In a similar way, Bayes factors for changes in covariance structure can be developed using modifications of the analysis given in Smith and Spiegelhalter (1981, sec. 17.3).

3. Changes of coefficients in regression models

3.1. Introduction

We shall consider the regression model

$$Y_t = x_t^T \beta^{(t)} + \varepsilon_t, \quad t = 1, \dots, n, \quad (22)$$

where, at time t , Y_t is the observation on the dependent variable, x_t is the column vector of observations on p regressor variables (including, possibly, a constant term), $\beta^{(t)}$ is the column vector of unknown regression coefficients, and ε_t is a disturbance term, assumed independently normally distributed for differing t , with zero mean and constant variance σ^2 .

The sequence defined by (22) will be said to have a change-point at r ($1 \leq r < n$) if

$$\beta^{(1)} = \dots = \beta^{(r)} = \beta, \quad \beta^{(r+1)} = \dots = \beta^{(n)} = \beta + \delta,$$

with unknown β , $\delta \neq 0$. We denote this model by M_r . The model of no change, $\delta = 0$, will be denoted by M_0 .

If we write $Y = (Y_1, \dots, Y_n)^T$ and denote a realization of Y by y , then adopting the notation

$$X_r^T = (x_1, \dots, x_r), \quad X_{(n-r)}^T = (x_{r+1}, \dots, x_n),$$

we see that, under M_r ,

$$Y \sim N_n(A_r \theta, \sigma^2 I_n), \quad (23)$$

where

$$A_r = \begin{pmatrix} X_r & 0 \\ X_{(n-r)} & X_{(n-r)} \end{pmatrix}, \quad \theta = \begin{pmatrix} \beta \\ \delta \end{pmatrix}, \quad (24)$$

whereas, under M_0 ,

$$Y \sim N_n(A_0\beta, \sigma^2 I_n), \quad (25)$$

where $A_0 = X_n$.

In (23)–(25), we have defined $p(y|M_r, \theta, \sigma)$ and $p(y|M_0, \beta, \sigma)$. To utilize (6)–(8) we require

$$p(y|M_r) = \iint p(y|M_r, \theta, \sigma) p(\theta, \sigma) d\theta d\sigma, \quad (26)$$

and

$$p(y|M_0) = \iint p(y|M_0, \beta, \sigma) p(\beta, \sigma) d\beta d\sigma, \quad (27)$$

we thus need to specify $p(\theta, \sigma)$ and $p(\beta, \sigma)$. This specification, and its relation to the problem of comparing alternative linear models, has been discussed extensively in the literature. Recent developments, together with extensive references, are presented in Smith and Spiegelhalter (1980) and Spiegelhalter and Smith (1982). In the next two sections, we shall consider two particular forms of prior specification.

3.2. A proper prior specification for δ

We consider first, following Smith (1980), the consequences of the specification

$$p(\theta, \sigma) = p(\theta|\sigma) p(\sigma), \quad (28)$$

where $p(\sigma) = c\sigma^{-1}$, and $p(\theta|\sigma)$ corresponds, for $1 \leq r \leq n$, to a normal distribution with mean θ_0 and covariance matrix $\sigma^2 V_0$, where

$$V_0 = \begin{bmatrix} V_{0\beta} & 0 \\ 0 & V_{0\delta} \end{bmatrix}, \quad \theta_0 = \begin{bmatrix} \beta_0 \\ \delta_0 \end{bmatrix}.$$

In the case of M_0 , we simply take $V_0 = V_{0\beta}$, $\theta_0 = \beta_0$. With this specification, it is easily verified [Smith (1980, sec. 3)] that if V_0^{-1} may be considered small in relation to $A_r^T A_r$, then

$$B_{r0} = \left(\frac{|X_n^T X_n|}{|V_{0\delta}| |X_r^T X_r| |X_{(n-r)}^T X_{(n-r)}|} \right)^{\frac{1}{2}} \left(1 + \frac{p}{(n-2p)} F_r \right)^{n/2}, \quad (29)$$

where F_r denotes the usual F -statistic for testing M_0 versus M_r .

General discussion:
conjugate prior distribution
change-point problem
and Cook (1980) and

3.3. Vague prior information

Returning to (28),

$$p(\theta, \sigma) =$$

$$= c$$

whereas, under M_0 ,

$$p(\beta, \sigma) = c$$

From (23)–(25), (30)
(1982) for a related

$$B_{r0} = c_\delta \left($$

Motivating a choice
less straightforward
argument is provided
the suggestion

$$c_\delta = (2p-1)$$

4. Changes of mean

4.1. General ideas

We assume that
infinite process which
(u, v) form

$$(1 - \phi_1 B$$

as in Box and Jenkins

General discussion of the comparison of linear models with informative conjugate prior distributions is given in Lempers (1971). Applications to change-point problems are given in Holbert and Broemeling (1977), Smith and Cook (1980) and Chin Choy and Broemeling (1980a, b).

3.3. Vague prior information for all parameters

Returning to (28), we now assume that, under M_r ,

$$p(\theta, \sigma) = p(\beta|\sigma)p(\delta|\sigma)p(\sigma) \quad (30)$$

$$= c_\beta (2\pi\sigma^2)^{-p/2} c_\delta (2\pi\sigma^2)^{-p/2} c_\sigma \sigma^{-1},$$

whereas, under M_0 ,

$$p(\beta, \sigma) = c_\beta (2\pi\sigma^2)^{-p/2} c_\sigma \sigma^{-1}. \quad (31)$$

From (23)–(25), (30) and (31), it easily follows [see Spiegelhalter and Smith (1982) for a related calculation] that

$$B_{r0} = c_\delta \left(\frac{|X_n^T X_n|}{|X_r^T X_r| |X_{(n-r)}^T X_{(n-r)}|} \right)^{\frac{1}{2}} \left(1 + \frac{p}{(n-2p)} F_r \right)^{n/2}. \quad (32)$$

Motivating a choice of c_δ by means of a calibrating 'thought experiment' is less straightforward than in the cases considered in section 2. A detailed argument is provided in Spiegelhalter and Smith (1982, sec. 4) and leads to the suggestion

$$c_\delta = (2p+1)^{p/2}. \quad (33)$$

4. Changes of mean in linear time series models

4.1. General ideas

We assume that observations $z = (z_1, \dots, z_n)^T$ are available from a doubly infinite process which may be modelled, for $t = \dots, -1, 0, 1, \dots$, by the ARMA (u, v) form

$$(1 - \phi_1 B - \dots - \phi_u B^u)(z_t - \mu_t) = (1 - \alpha_1 B - \dots - \alpha_v B^v)\varepsilon_t, \quad (34)$$

as in Box and Jenkins (1970). We then denote by M_r the model which

(25)

(26)

(27)

cification, and its relation
odels, has been discussed
, together with extensive
(1980) and Spiegelhalter
ll consider two particular

ie consequences of the

(28)

< n, to a normal distri-
here

With this specification, it
ay be considered small in

$$\left(\frac{2}{2p} F_r \right)^{n/2}, \quad (29)$$

versus M_r .

assumes that

$$\left. \begin{aligned} \mu_t &= \beta, & t &= 1, \dots, r \\ \mu_t &= \beta + \delta, & t &= r+1, \dots, n \end{aligned} \right\}, \quad (35)$$

and by M_0 the model which assumes that

$$\mu_t = \beta, \quad t = 1, \dots, n. \quad (36)$$

Using ideas from Box and Tiao (1965) and Glass et al. (1975), we seek a linear transformation from $z = (z_1, \dots, z_n)^T$ to $y = (y_1, \dots, y_n)^T$, such that, under M_r , the random vector Y , corresponding to y , can be described by

$$Y \sim N_n(A_r \theta, \sigma^2 I_n), \quad (37)$$

where $A_r = A_r(\phi, \alpha)$ is an $n \times 2$ matrix, whose entries are functions of r and $(\phi^T, \alpha^T) = (\phi_1, \dots, \phi_n, \alpha_1, \dots, \alpha_n)$, and $\theta = (\beta, \delta)^T$. Under M_0 , we require

$$Y \sim N_n(A_0 \beta, \sigma^2 I_n), \quad (38)$$

where $A_0 = A_0(\phi, \alpha)$ is $n \times 1$ and given by the first column of A_r (corresponding to $\delta = 0$).

Before turning to a detailed discussion of the form of such a transformation, we note the following important relationship with the work of section 3. If $p(\theta, \sigma)$, $p(\beta, \sigma)$ are defined either as in (18), or by (30), (31), with $p = 1$, we can calculate

$$p(z | M_r, \phi, \alpha) = \iint p(z | M_r, \theta, \sigma, \phi, \alpha) p(\theta, \sigma) d\theta d\sigma, \quad (39)$$

and

$$p(z | M_0, \phi, \alpha) = \iint p(z | M_0, \beta, \sigma, \phi, \alpha) p(\beta, \sigma) d\beta d\sigma, \quad (40)$$

where the first terms in the integrals are defined by (37) and (38) multiplied by the Jacobian of the transformation, $J(\phi, \alpha)$, say. If we then integrate (39) and (40) with respect to a prior specification $p(\phi, \alpha)$ for the ARMA parameters, thus obtaining $p(z | M_r)$, $p(z | M_0)$, the ratio of these latter terms gives $B_{r,0}$. The resulting form is

$$B_{r,0} = K_s \frac{\iint J(\phi, \alpha) |A_r^T(\phi, \alpha) A_r(\phi, \alpha)|^{-1/2} [S_r(\phi, \alpha)]^{-n/2} p(\phi, \alpha) d\phi d\alpha}{\iint J(\phi, \alpha) |A_0^T(\phi, \alpha) A_0(\phi, \alpha)|^{-1/2} [S_0(\phi, \alpha)]^{-n/2} p(\phi, \alpha) d\phi d\alpha}, \quad (41)$$

where $K_s = V_{00}^{-1/2}$ [if $p(\theta, \sigma)$ defined by (30)], and S_r squares resulting from least squares (37) and (38).

In most applications, u the integrals arising in (4) covering the admissible region of ϕ, α .

The general approach follows.

If $s = \max(u, v)$, the first independent y_i 's by definition

$$y_1 = \lambda_1 z_1$$

$$y_2 = \lambda_2(z_2 + \pi_2 z_1)$$

$$\vdots$$

$$y_s = \lambda_s(z_s + \pi_{s2} z_2 + \dots + \pi_{s1} z_1)$$

where the λ_i 's are chosen 1 are chosen successively in $s < i \leq n$, we define

$$y_i = \lambda_i(z_i - \phi_{i1} z_1 - \dots - \phi_{i,i-1} z_{i-1})$$

where the successive choice and independence, respectively.

For AR(u) models, all ARMA(u, v) models with $u+v \leq 2$ are defined recursively, at least terms. In the following sections cases with $u+v \leq 2$. If $u+v$ first $u+v$ y_i 's. An alternative T^{-1} of the transformation relation $TT^T = \sigma^{-2} \Gamma$, where z_{u+v} . It will be seen that problem of finding the example, Newbold (1974) these authors invert the observations, the method

where $K_s = V_{0s}^{-1}$ [if $p(\theta, \sigma)$ is defined by (28)], or $K_s = c_s = \sqrt{3}$ [if $p(\theta, \sigma)$ is defined by (30)], and $S_r(\phi, \alpha)$, $S_o(\phi, \alpha)$ denote the usual residual sum of squares resulting from least squares fitting of the linear models defined by (37) and (38).

In most applications, u and v are small and an adequate approximation to the integrals arising in (41) is obtained by summation over a suitable grid covering the admissible region (with respect to stationarity and invertibility) of ϕ, α .

The general approach to deriving an appropriate transformation is as follows.

If $s = \max(u, v)$, the first s z_i 's must be transformed directly to a set of s independent y_i 's by defining

$$\left. \begin{aligned} y_1 &= \lambda_1 z_1 \\ y_2 &= \lambda_2 (z_2 + \pi_{22} z_1) \\ &\vdots \\ y_s &= \lambda_s (z_s + \pi_{s2} z_{s-1} + \cdots + \pi_{ss} z_1) \end{aligned} \right\}, \quad (42)$$

where the λ_i 's are chosen to reduce the variance of the y_i 's to σ^2 , and the π_{ij} 's are chosen successively in such a way that the y_i 's are independent. For $s < i \leq n$, we define

$$y_i = \lambda_i (z_i - \phi_1 z_{i-1} - \cdots - \phi_u z_{i-u} - \tau_{i1} y_{i-1} - \cdots - \tau_{iv} y_{i-v}), \quad (43)$$

where the successive choices of the λ_i 's and the τ_{ij} 's ensures homoscedasticity and independence, respectively.

For AR(u) models, all the τ_{ij} terms are 0 and $\lambda_i = 1$ for $s < i \leq n$. For ARMA(u, v) models with $u > 0, v > 0$, both the τ_{ij} and λ_i terms are eventually defined recursively, at least from $i = p + q + 1$. MA(v) models lose the ϕ_{p+i-j} terms. In the following sections, we provide a detailed development of all cases with $u + v \leq 2$. If $u + v$ is large, it may be difficult to identify directly the first $u + v$ y_i 's. An alternative approach is to find the lower triangular matrix T^{-1} of the transformation from (z_1, \dots, z_{n+v}) to (y_1, \dots, y_{n+v}) by using the relation $TT^T = \sigma^{-2} \Gamma$, where Γ is the variance-covariance matrix of z_1, \dots, z_{n+v} . It will be seen that the techniques used here are closely related to the problem of finding the exact likelihood for ARMA processes [see, for example, Newbold (1974) and Galbraith and Galbraith (1974), but whereas these authors invert the variance-covariance matrix of a sequence of observations, the method proposed here effectively diagonalises it].

(35)

(36)

ss et al. (1975), we seek a
..., y_n), such that, under
described by

(37)

ss are functions of r and
in M_0 , we require

(38)

column of A , (correspond-

form of such a trans-
relationship with the work of
(18), or by (30), (31), with

ld $d\sigma$, (39)

ld $d\sigma$, (40)

(37) and (38) multiplied
If we then integrate (39)
 $p(\phi, \alpha)$ for the ARMA
tio of these latter terms

$$\frac{[\alpha]^{-n/2} p(\phi, \alpha) d\phi d\alpha}{[\alpha]^{-n/2} p(\phi, \alpha) d\phi d\alpha}, \quad (41)$$

4.2. The AR(1) case

The required transformation is defined by

$$\left. \begin{aligned} y_1 &= (1 - \phi_1^2)^{\frac{1}{2}} z_1 \\ y_i &= z_i - \phi_1 z_{i-1}, \quad i = 2, \dots, n \end{aligned} \right\}, \quad (44)$$

with Jacobian $(1 - \phi_1^2)^{\frac{1}{2}}$. The A_r matrix has the form

$$A_r^T = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{bmatrix}, \quad (45)$$

where

$$\left. \begin{aligned} a_1 &= (1 - \phi_1^2)^{\frac{1}{2}}, \quad a_2 = a_3 = \dots = a_n = (1 - \phi_1) \\ b_1 &= b_2 = \dots = b_r = 0, \quad b_{r+1} = 1, \quad b_{r+2} = \dots = b_n = (1 - \phi_1) \end{aligned} \right\}. \quad (46)$$

4.3. The AR(2) case

The required transformation is defined by

$$\left. \begin{aligned} y_1 &= (\gamma_0/\sigma^2)^{-\frac{1}{2}} z_1 \\ y_2 &= [(\gamma_0/\sigma^2)(1 - \rho_1^2)]^{-\frac{1}{2}} (z_2 - \rho_1 z_1) \\ y_i &= z_i - \phi_1 z_{i-1} - \phi_2 z_{i-2}, \quad i = 3, \dots, n \end{aligned} \right\}, \quad (47)$$

where

$$(\gamma_0/\sigma^2) = (1 - \phi_2)[(1 + \phi_2)\{(1 - \phi_2)^2 - \phi_1^2\}]^{-1}, \quad \rho_1 = \phi_1(1 - \phi_2)^{-1},$$

with Jacobian $(\gamma_0/\sigma^2)^{-\frac{1}{2}}(1 - \rho_1^2)^{-\frac{1}{2}}$. The A_r matrix, (45), is defined by

$$\left. \begin{aligned} a_1 &= (\gamma_0/\sigma^2)^{-\frac{1}{2}}, \quad a_2 = [(\gamma_0/\sigma^2)(1 - \rho_1^2)]^{-\frac{1}{2}}(1 - \rho_1) \\ a_3 &= \dots = a_n = 1 - \phi_1 - \phi_2 \\ b_1 &= \dots = b_r = 0, \quad b_{r+1} = 1, \quad b_{r+2} = 1 - \phi_1 \\ b_{r+3} &= \dots = b_n = 1 - \phi_1 - \phi_2 \end{aligned} \right\}. \quad (48)$$

4.4. The MA(1) case

The required transform

$$\begin{aligned} y_1 &= \lambda_1 z_1 \\ y_i &= \lambda_i(z_i - \tau_i) \end{aligned}$$

where

$$\begin{aligned} \lambda_1 &= (\gamma_0/\sigma^2)^{-\frac{1}{2}} \\ \lambda_i &= (1 + \alpha_1^2 - \end{aligned}$$

The Jacobian of the tra:

$$J(\phi, \alpha) = \lambda_1 \lambda$$

and the elements of the

$$a_1 = \lambda_1, \quad a_i =$$

$$b_1 = \dots = b_r =$$

We can solve (52) to ob

$$a_i = (1 - \alpha_i)^{-}$$

$$b_{r+i} = (1 - \alpha_i$$

and some simplification
 $i \rightarrow \infty, \lambda_i \rightarrow 1, a_i \rightarrow (1 -$
 $J(\phi, \alpha) \rightarrow (1 - \alpha_i^2)^{\frac{1}{2}}.$

4.5. The MA(2) case

The required transform

$$\begin{aligned} y_1 &= \lambda_1 z_1 \\ y_2 &= \lambda_2(z_2 + \tau \\ y_i &= \lambda_i(z_i + \tau_i) \end{aligned}$$

4.4. The MA(1) case

The required transformation is defined by

$$\left. \begin{aligned} y_1 &= \lambda_1 z_1 \\ y_i &= \lambda_i (z_i - \tau_i y_{i-1}), \quad i=2, \dots, n \end{aligned} \right\}, \quad (49)$$

where

$$\left. \begin{aligned} \lambda_1 &= (\gamma_0/\sigma^2)^{-\frac{1}{2}} = (1 + \alpha_1^2)^{-\frac{1}{2}}, \quad \tau_i = -\alpha_1 \lambda_{i-1}, \\ \lambda_i &= (1 + \alpha_1^2 - \tau_i^2)^{-\frac{1}{2}} = (1 - \alpha_1^2)^{\frac{1}{2}} (1 - \alpha_1^{2(i+1)})^{-\frac{1}{2}}, \end{aligned} \right\} \quad i=2, \dots, n. \quad (50)$$

The Jacobian of the transformation is given by

$$J(\phi, \alpha) = \lambda_1 \lambda_2 \dots \lambda_n = (1 - \alpha_1^2)^{\frac{n}{2}} (1 - \alpha_1^{2(n+1)})^{-\frac{1}{2}}, \quad (51)$$

and the elements of the A_r matrix, (45), are defined by

$$\left. \begin{aligned} a_1 &= \lambda_1, \quad a_i = \lambda_i (1 - \tau_i a_{i-1}), \quad i=2, \dots, n \\ b_1 &= \dots = b_r = 0, \quad b_i = \lambda_i (1 - \tau_i b_{i-1}), \quad i=r+1, \dots, n \end{aligned} \right\}. \quad (52)$$

We can solve (52) to obtain explicit forms,

$$\left. \begin{aligned} a_i &= (1 - \alpha_1)^{-1} (1 - \alpha_1^i)^{\frac{1}{2}} (1 - \alpha_1^{i+1})^{\frac{1}{2}} (1 + \alpha_1^i)^{-\frac{1}{2}} (1 + \alpha_1^{i+1})^{-\frac{1}{2}}, \\ b_{r+i} &= (1 - \alpha_1)^{-1} (1 - \alpha_1^i) (1 - \alpha_1^{2r+i+1}) (1 - \alpha_1^{2(r+i)})^{-\frac{1}{2}} (1 - \alpha_1^{2(r+i+1)})^{-\frac{1}{2}}, \end{aligned} \right\} \quad i=1, \dots, n, \quad (47)$$

and some simplifications are possible for large i and n . For example, as $i \rightarrow \infty$, $\lambda_i \rightarrow 1$, $a_i \rightarrow (1 - \alpha_1)^{-1}$, $b_{r+i} \rightarrow (1 - \alpha_1)^{-1}$, $\tau_i \rightarrow -\alpha_1$, and, as $n \rightarrow \infty$, $J(\phi, \alpha) \rightarrow (1 - \alpha_1^2)^{\frac{n}{2}}$.

4.5. The MA(2) case

The required transformation is defined by

$$\left. \begin{aligned} y_1 &= \lambda_1 z_1 \\ y_2 &= \lambda_2 (z_2 + \tau_2 y_1) \\ y_i &= \lambda_i (z_i + \tau_i y_{i-1} + \pi_i y_{i-2}), \quad i=3, \dots, n \end{aligned} \right\}, \quad (53)$$

5), is defined by

$$\left. \begin{aligned} (1 - \rho_1) \\ \phi_1 \end{aligned} \right\}. \quad (48)$$

$$\rho_1 = \phi_1 (1 - \phi_2)^{-1},$$

where

$$\left. \begin{aligned} \lambda_1 &= (\gamma_0/\sigma^2)^{-\frac{1}{2}}, \quad \tau_2 = -\gamma_1\lambda_1/\sigma^2, \quad \lambda_2 = [(\gamma_0/\sigma^2)(1-\tau_2^2)]^{-\frac{1}{2}} \\ \gamma_0 &= (1+\alpha_1^2+\alpha_2^2)\sigma^2, \quad \gamma_1 = (-\alpha_1+\alpha_1\alpha_2)\sigma^2, \quad \gamma_2 = -\alpha_2\sigma^2 \end{aligned} \right\} \quad (54)$$

and

$$\left. \begin{aligned} \lambda_i &= (1+\alpha_1^2+\alpha_2^2-\tau_i^2-\pi_i^2)^{-\frac{1}{2}} \\ \pi_i &= -\lambda_{i-2}\gamma_2/\sigma^2 \\ \tau_i &= -\lambda_{i-1}(\gamma_1+\tau_{i-1}\lambda_{i-2}\gamma_2)/\sigma^2 \end{aligned} \right\} \quad (55)$$

The Jacobian of the transformation is given by $\lambda_1 \lambda_2 \dots \lambda_n$ and the elements of A_r are defined by

$$\left. \begin{aligned} a_1 &= \lambda_1, \quad a_2 = \lambda_2(1+\tau_2 a_1) \\ a_i &= \lambda_i(1+\tau_i a_{i-1} + \pi_i a_{i-2}), \quad i=3, \dots, n \\ b_1 &= \dots = b_r = 0 \\ b_{r+i} &= \lambda_{r+i}(1+\tau_{r+i} b_{r+i-1} + \pi_{r+i} b_{r+i-2}), \quad i=1, \dots, n-r \end{aligned} \right\} \quad (56)$$

For large values of i , $\lambda_i \rightarrow 1$, $\tau_i \rightarrow \alpha_1$, $\pi_i \rightarrow \alpha_2$, $a_i \rightarrow (1-\alpha_1-\alpha_2)^{-1}$, $b_i \rightarrow (1-\alpha_1-\alpha_2)^{-1}$.

4.6. The ARMA(1,1) case

The required transformation is defined by

$$\left. \begin{aligned} y_1 &= \lambda_1 z_1 \\ y_i &= \lambda_i(z_i - \phi_1 z_{i-1} - \tau_i y_{i-1}), \quad i=2, \dots, n \end{aligned} \right\} \quad (57)$$

where

$$\left. \begin{aligned} \lambda_1 &= (\gamma_0/\sigma^2)^{-\frac{1}{2}} = (\alpha_1^2 - 2\phi_1\alpha_1 + 1)^{-\frac{1}{2}}(1-\phi_1^2)^{\frac{1}{2}} \\ \tau_i &= -\alpha_1\lambda_{i-1}, \quad \lambda_i = (1+\alpha_1^2-\tau_i^2)^{-\frac{1}{2}}, \quad i=2, \dots, n \end{aligned} \right\} \quad (58)$$

Explicitly, we have

$$\lambda_i = (C - D\alpha_1^{2(i-1)})^{\frac{1}{2}}(C - D\alpha_1^2)^{-\frac{1}{2}}, \quad (59)$$

where

$$C = (1-\phi_1\alpha_1)^2$$

and the Jacobian is given

$$J(\phi, \alpha) = \lambda_1 \lambda_2$$

The elements of the A_r

$$a_1 = \lambda_1, \quad a_i =$$

$$b_1 = \dots = b_r =$$

$$b_i = \lambda_i(1-\phi_1-$$

leading to the explicit for

$$a_i = \frac{(1-\phi_1)(1-$$

$$b_{r+i} = \frac{(1-\phi_1)}{($$

As $i \rightarrow \infty$, $\tau_i \rightarrow -\alpha_1$, $a_i \rightarrow$
 $n \rightarrow \infty$,

$$J(\phi, \alpha) \rightarrow (1-\phi$$

References

- Atkinson, A.C., 1978, Posterior 48.
- Box, G.E.P. and G.M. Jenkins, San Francisco, CA).
- Box, G.E.P. and G.C. Tiao, Biometrika 52, 181-192.
- Broemeling, L.D., 1972, Bayesian variables, Metron XXX, 1-1.
- Broemeling, L.D., 1974, Bayesian Communications in Statistic
- Chin Choy, J.H. and L.D. Bro model, Technometrics 22, 71

where

$$C = (1 - \phi_1 \alpha_1)^2, \quad D = (\alpha_1 - \phi_1)^2, \quad (60)$$

and the Jacobian is given by

$$J(\phi, \alpha) = \lambda_1 \lambda_2 \dots \lambda_n = (1 - \phi_1^2)(1 - \alpha_1^2)(C - D\alpha_1^2)^{-\frac{1}{2}}. \quad (61)$$

The elements of the A_r matrix are given by

$$\left. \begin{aligned} a_i &= \lambda_i, \quad a_i = \lambda_i(1 - \phi_1 - \tau_i a_{i-1}), \quad i = 2, \dots, n \\ b_1 &= \dots = b_r = 0, \quad b_{r+1} = \lambda_{r+1} \\ b_i &= \lambda_i(1 - \phi_1 - \tau_i b_{i-1}), \quad i = r+2, \dots, n \end{aligned} \right\}, \quad (62)$$

leading to the explicit forms

$$\left. \begin{aligned} a_i &= \frac{(1 - \phi_1)(1 - \alpha_1^{i-1})(C - D\alpha_1^i) + \alpha_1^{i-1}(1 - \alpha_1)(1 - \alpha_1^2)(1 - \phi_1^2)}{(1 - \alpha_1)(C - D\alpha_1^{2(i-1)})^{\frac{1}{2}}(C - D\alpha_1^{2i})^{\frac{1}{2}}}, \\ b_{r+i} &= \frac{(1 - \phi_1)(1 - \alpha_1^{i-1})(C - D\alpha_1^{2r+i}) + \alpha_1^{i-1}(1 - \alpha_1)(C - D\alpha_1^{2r})}{(1 - \alpha_1)(C - D\alpha_1^{2(r+i-1)})^{\frac{1}{2}}(C - D\alpha_1^{2(r+i)})^{\frac{1}{2}}}, \end{aligned} \right\} \quad i = 1, \dots, n \quad (63)$$

As $i \rightarrow \infty$, $\tau_i \rightarrow -\alpha_1$, $a_i \rightarrow (1 - \phi_1)(1 - \alpha_1)^{-1}$, $b_{r+i} \rightarrow (1 - \phi_1)(1 - \alpha_1)^{-1}$ and, as $n \rightarrow \infty$,

$$J(\phi, \alpha) \rightarrow (1 - \phi_1^2)^{\frac{1}{2}}(1 - \alpha_1^2)^{\frac{1}{2}}(1 - \phi_1 \alpha_1)^{-1}. \quad (57)$$

References

- Atkinson, A.C., 1978, Posterior probabilities for choosing a regression model, *Biometrika* 65, 39-48.
- Box, G.E.P. and G.M. Jenkins, 1970, *Time series analysis: Forecasting and control* (Holden-Day, San Francisco, CA).
- Box, G.E.P. and G.C. Tiao, 1965, A change in the level of a non-stationary time-series, *Biometrika* 52, 181-192.
- Broemeling, L.D., 1972, Bayesian procedures for detecting a change in a sequence of random variables, *Metron* XXX, 1-14.
- Broemeling, L.D., 1974, Bayesian inferences about a changing sequence of random variables, *Communications in Statistics* 3, 243-255.
- Chin Choy, J.H. and L.D. Broemeling, 1980a, Some Bayesian inferences for a changing linear model, *Technometrics* 22, 71-78.

- Chin Choy, J.H. and L.D. Broemeling, 1980b, Detecting the change in a linear model, Technical report no. 14 (Department of Statistics, Oklahoma State University, Stillwater, OK).
- Ferreira, P.E., 1975, A Bayesian analysis of a switching regression model: A known number of regimes, *Journal of the American Statistical Association* 70, 370-374.
- Galbraith, R.F. and J.I. Galbraith, 1974, On the inverses of some patterned matrices arising in the theory of linear time-series, *Journal of Appl. Prob.* 11, 63-71.
- Glass, G.V., J.M. Gottman and V.L. Willson, 1975, Design and analysis, of time-series experiments (Colorado Associated University Press, Boulder, CO).
- Harrison, P.J. and C.F. Stevens, 1976, Bayesian forecasting (with discussion), *Journal of the Royal Statistical Society B38*, 205-247.
- Holbert, D. and L.D. Broemeling, 1977, Bayesian inference related to shifting sequences and two-phase regression, *Communications in Statistics* 6.
- Jeffreys, H., 1936/1961, *Theory of probability* (Oxford University Press, Oxford).
- Lee, A.F.S. and S.M. Heghinian, 1977, A shift of the mean level in a sequence of independent normal random variables - A Bayesian approach, *Technometrics* 19, 503-506.
- Lempers, F.B., 1971, *Posterior probabilities of alternative linear models* (Rotterdam University Press, Rotterdam).
- Menzefricke, U., 1981, A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point, *Applied Statistics* 30, 141-146.
- Newbold, P., 1974, The exact likelihood function for a mixed autoregressive-moving-average process, *Biometrika* 61, 423-426.
- Smith, A.F.M., 1975, A Bayesian approach to inference about a change-point in a sequence of random variables, *Biometrika* 62, 407-416.
- Smith, A.F.M., 1977, A Bayesian analysis of some time-varying models, in: J.R. Barra et al., *Recent developments in statistics* (North-Holland, Amsterdam) 257-267.
- Smith, A.F.M., 1980, Change-point problems: Approaches and applications, in: J.M. Bernardo et al., eds., *Bayesian statistics* (Valencia, University Press, Valencia) 83-98.
- Smith, A.F.M. and D.G. Cook, 1980, Switching straight lines: A Bayesian analysis of some renal transplant data, *Applied Statistics* 29, 180-184.
- Smith, A.F.M. and U.E. Makov, 1980, Bayesian detection and estimation of jumps in linear systems, in: O.L.R. Jacobs et al., eds., *The analysis and optimization of stochastic systems* (Academic Press, London) 333-346.
- Smith, A.F.M. and D.J. Spiegelhalter, 1980, Bayes factors and choice criteria for linear models, *Journal of the Royal Statistical Society B42*, 213-220.
- Smith, A.F.M. and D.J. Spiegelhalter, 1981, Bayesian approaches to multivariate structure, in: V. Barnett, ed., *Interpreting multivariate data* (Wiley, Chichester) 335-348.
- Smith, A.F.M. and M. West, 1982, Monitoring renal transplants: An application of the multi-process Kalman filter, Submitted for publication.
- Spiegelhalter, D.J. and A.F.M. Smith, 1982, Bayes factors for linear and log-linear models with vague prior information, *Journal of the Royal Statistical Society B44*, forthcoming.

BAYESIAN PARAMETER

1. Introduction

The gamma distribution

$$f(x; \theta) = \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x}$$

where k is a known widely used as a model for practical examples in

Under certain conditions occurs at least once is (1). In other words x_1, \dots, x_n of independent

$$x_i \sim f(x; \theta_i)$$

$$\sim f(x; \theta)$$

where f corresponds to λ is the parameter that take values between therefore $\theta_1 = \theta_2$.

Under the previous first is the detection change point and t change has occurred

Several authors problems for the n Bacon and Watts (Holbert and Broemeling)

0165-7410/82/0000-

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.